RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/536,772
Source: PKT
Date Processed by STIC: 6/13/05

ENTERED

CRF Errors Edited by the STIC Systems Branch

Serial Number: 10/536,772 CRF Edit Date: 6/13/05
Edited by: JCL

___ Realigned nucleic acid/amino acid numbers/text in cases where the sequence text "wrapped" to the next line

___ Corrected the SEQ ID NO. Sequence numbers edited were:

_____

___ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

_____

✓ Deleted: ✓ invalid beginning/end-of-file text ; ___ page numbers

___ Inserted mandatory headings/numeric identifiers, specifically:

_____

___ Moved responses to same line as heading/numeric identifier, specifically:

_____

___ Other:

_____

_____

_____

Revised 09/09/2003

Raw Sequence Listing before editing,
for reference only

PCT

RAW SEQUENCE LISTING DATE: 06/13/2005
PATENT APPLICATION: **US/10/536,772** TIME: 15:05:39

Input Set : **A:\pto.kd.txt**
Output Set: **N:\CRF4\06132005\J536772.raw**

```
      4 <110> APPLICANT: SHANGHAI CANCER INSTITUTE
      6 <120> TITLE OF INVENTION: A HUMAN TUMOR-ASSOCIATED GENE CT120 ON CHROMOSOME 17P 13.3
REGION AND
      7       PROTEIN
      8       ENCODED BY IT
     10 <130> FILE REFERENCE: 024832pc
C--> 12 <140> CURRENT APPLICATION NUMBER: US/10/536,772
C--> 12 <141> CURRENT FILING DATE: 2005-05-26
     12 <150> PRIOR APPLICATION NUMBER: CN 02150730.9
     13 <151> PRIOR FILING DATE: 2002-11-27
     15 <160> NUMBER OF SEQ ID NOS: 13
     17 <170> SOFTWARE: PatentIn version 3.1
     19 <210> SEQ ID NO: 1
     20 <211> LENGTH: 2145
     21 <212> TYPE: DNA
     22 <213> ORGANISM: Homo sapiens
     24 <220> FEATURE:
     25 <221> NAME/KEY: CDS
     26 <222> LOCATION: (91)..(861)
     27 <223> OTHER INFORMATION:
W--> 29 <400> 1
     30 cggagggttg aaatcgcgcg gccgggccgg ggcgcgccga gccgaaccca gccacgcggc     60
     31 gccagcgagg cggccggacc cgcagccccg atg ctg ctg acg ctg gcc ggg ggc     114
     32                                  Met Leu Leu Thr Leu Ala Gly Gly
     33                                  1               5
     34 gcg ctc ttc ttc ccg ggg ctc ttc gcg ctc tgc acc tgg gcg ctg cgc     162
     35 Ala Leu Phe Phe Pro Gly Leu Phe Ala Leu Cys Thr Trp Ala Leu Arg
     36     10                  15                  20
     37 cac tcc cag ccc gga tgg agc cgc acc gac tgc gtg atg atc agc acc     210
     38 His Ser Gln Pro Gly Trp Ser Arg Thr Asp Cys Val Met Ile Ser Thr
     39 25                  30                  35                  40
     40 agg ctg gtt tcc tcg gtg cac gcc gtg ctg gcc acc ggc tcg ggg atc     258
     41 Arg Leu Val Ser Ser Val His Ala Val Leu Ala Thr Gly Ser Gly Ile
     42                 45                  50                  55
     43 gtc atc att cgc tcc tgc gac gac gtg atc acc ggc agg cac tgg ctt     306
     44 Val Ile Ile Arg Ser Cys Asp Asp Val Ile Thr Gly Arg His Trp Leu
     45             60                  65                  70
     46 gcc cgg gaa tat gtg tgg ttt ctg att cca tac atg atc tat gac tcg     354
     47 Ala Arg Glu Tyr Val Trp Phe Leu Ile Pro Tyr Met Ile Tyr Asp Ser
     48         75                  80                  85
     49 tac gcc atg tac ctc tgt gaa tgg tgc cga acc aga gac cag aac cgt     402
     50 Tyr Ala Met Tyr Leu Cys Glu Trp Cys Arg Thr Arg Asp Gln Asn Arg
     51     90                  95                  100
     52 gcg ccc tcc ctc act ctt cga aac ttc cta agt cga aac cgc ctc atg     450
```

RAW SEQUENCE LISTING DATE: 06/13/2005
PATENT APPLICATION: US/10/536,772 TIME: 15:05:39

Input Set : A:\pto.kd.txt
Output Set: N:\CRF4\06132005\J536772.raw

```
53 Ala Pro Ser Leu Thr Leu Arg Asn Phe Leu Ser Arg Asn Arg Leu Met
54 105                 110                 115                 120
55 atc aca cat cat gcg gtc att ctc ctt gtc ctt gtg cca gtc gca cag      498
56 Ile Thr His His Ala Val Ile Leu Leu Val Leu Val Pro Val Ala Gln
57                 125                 130                 135
58 agg ctc cgg gga gac ctt ggg gac ttc ttt gtc ggc tgc atc ttc acg      546
59 Arg Leu Arg Gly Asp Leu Gly Asp Phe Phe Val Gly Cys Ile Phe Thr
60             140                 145                 150
61 gca gaa ctg agc act ccg ttt gtg tcg ctg ggc agg gtt ctg att cag      594
62 Ala Glu Leu Ser Thr Pro Phe Val Ser Leu Gly Arg Val Leu Ile Gln
63         155                 160                 165
64 cta aag cag cag cac acc ctt ctg tac aag gtg aat gga atc ctc acg      642
65 Leu Lys Gln Gln His Thr Leu Leu Tyr Lys Val Asn Gly Ile Leu Thr
66     170                 175                 180
67 ctg gcc acc ttc ctt tcc tgc cgg atc ctt ctc ttc ccc ttc atg tac      690
68 Leu Ala Thr Phe Leu Ser Cys Arg Ile Leu Leu Phe Pro Phe Met Tyr
69 185                 190                 195                 200
70 tgg tcc tat ggc cgc cag cag gga cta agc ctg ctc caa gta ccc ttc      738
71 Trp Ser Tyr Gly Arg Gln Gln Gly Leu Ser Leu Leu Gln Val Pro Phe
72                 205                 210                 215
73 agc atc cca ttc tac tgc aac gtg gcc aat gcc ttc ctc gta gct cct      786
74 Ser Ile Pro Phe Tyr Cys Asn Val Ala Asn Ala Phe Leu Val Ala Pro
75             220                 225                 230
76 cag atc tac tgg ttc tgt ctg ctg tgc agg aag gca gtc cgg ctc ttt      834
77 Gln Ile Tyr Trp Phe Cys Leu Leu Cys Arg Lys Ala Val Arg Leu Phe
78         235                 240                 245
79 gac act ccc caa gcc aaa aag gat ggc taaatgctcc tgggagtcag            881
80 Asp Thr Pro Gln Ala Lys Lys Asp Gly
81     250                 255
82 gcgcagcctc acaccagctg cctcctccac tcagcattcc atggaccaaa ttgtgccctg    941
83 ggtagcctca gactttgggt attgataagc cgatggattt gagtttttct aaagaatatt   1001
84 catattacct ccttttttcta acttgcccta tttgcaaacg cacttttgta gtaacaacta   1061
85 ttgggtcctg tcagacctcc acggacagca aagtggtttt aatgcaagcc caaggatcct   1121
86 tcttaaggtc ttatctcaag agctctggga ggtggaagca tggggtggga tcggtggacc   1181
87 agggtggtaa gtgtctgcac atctgcctgt ccctgtatca gcggctaccc accttccaaa   1241
88 ccactcagga cagtacccgt ggcactgggc ccgcagaagc aagggatgac ttggttcttg   1301
89 gaagtaatgt cgtcttgtga cattggcctg ggacaatcat tgtgggtagg tagttattga   1361
90 tcgtttacta gataacccat tggttctttg cctcatcctc tcatccatgg gtcagagttg   1421
91 aattcttatg tctatagact tccaatcaga agtctcactg gtggggctgg gggtgggggc   1481
92 aggcaggagg catggatggg aacctgagta ggtagtgtgg ccaagagatc agcacaacct   1541
93 ttgcaggctg acttgctaag tctgacagtg acaaacttgt gagcttactg cagtcagtca   1601
94 cagaggctgt tctttttcac acacccctte atgcccggct ttccccatat ccacatgcag   1661
95 agggcgagct cataaaacta cagggaagcg tgaaatgatg gctttggtag ctgtttactg   1721
96 ggtaacccca ctgtgacact gtccttttca tgtgatgtgg aaacctactt ctgtcctcca   1781
97 aaccatgaaa tgtgtcatct agactgcaga gtactcgagt gctttgcctc ccgatatgcc   1841
98 agagcttgtg gtccaaagcc cattcctgtg tgtccgtcct gccatttagc cacagaaggc   1901
99 tgcggagtga ggcggcagct agcctggcca gtggctgtcc cgtggaccga cacctgcgcc   1961
100 cccttctgca agcaggattt tctggtgcca acactcattc atcattcccg atcaactagg   2021
101 atgaatttaa gactgtgcta ccatgtgttc tcaagtggta gtttaaaaag tggattttta   2081
```

RAW SEQUENCE LISTING DATE: 06/13/2005
PATENT APPLICATION: **US/10/536,772** TIME: 15:05:39

Input Set : **A:\pto.kd.txt**
Output Set: **N:\CRF4\06132005\J536772.raw**

aagtgccttt caattgtctg tgaacgtcta aaggactgat ttgtctcaaa aaaaaaaaaa 2141
aaaa 2145
<210> SEQ ID NO: 2
<211> LENGTH: 257
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 2
Met Leu Leu Thr Leu Ala Gly Gly Ala Leu Phe Phe Pro Gly Leu Phe
1 5 10 15
Ala Leu Cys Thr Trp Ala Leu Arg His Ser Gln Pro Gly Trp Ser Arg
20 25 30
Thr Asp Cys Val Met Ile Ser Thr Arg Leu Val Ser Ser Val His Ala
35 40 45
Val Leu Ala Thr Gly Ser Gly Ile Val Ile Ile Arg Ser Cys Asp Asp
50 55 60
Val Ile Thr Gly Arg His Trp Leu Ala Arg Glu Tyr Val Trp Phe Leu
65 70 75 80
Ile Pro Tyr Met Ile Tyr Asp Ser Tyr Ala Met Tyr Leu Cys Glu Trp
85 90 95
Cys Arg Thr Arg Asp Gln Asn Arg Ala Pro Ser Leu Thr Leu Arg Asn
100 105 110
Phe Leu Ser Arg Asn Arg Leu Met Ile Thr His His Ala Val Ile Leu
115 120 125
Leu Val Leu Val Pro Val Ala Gln Arg Leu Arg Gly Asp Leu Gly Asp
130 135 140
Phe Phe Val Gly Cys Ile Phe Thr Ala Glu Leu Ser Thr Pro Phe Val
145 150 155 160
Ser Leu Gly Arg Val Leu Ile Gln Leu Lys Gln Gln His Thr Leu Leu
165 170 175
Tyr Lys Val Asn Gly Ile Leu Thr Leu Ala Thr Phe Leu Ser Cys Arg
180 185 190
Ile Leu Leu Phe Pro Phe Met Tyr Trp Ser Tyr Gly Arg Gln Gln Gly
195 200 205
Leu Ser Leu Leu Gln Val Pro Phe Ser Ile Pro Phe Tyr Cys Asn Val
210 215 220
Ala Asn Ala Phe Leu Val Ala Pro Gln Ile Tyr Trp Phe Cys Leu Leu
225 230 235 240
Cys Arg Lys Ala Val Arg Leu Phe Asp Thr Pro Gln Ala Lys Lys Asp
245 250 255
Gly
<210> SEQ ID NO: 3
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(25)
<223> OTHER INFORMATION: primer
<400> SEQUENCE: 3

RAW SEQUENCE LISTING DATE: 06/13/2005
PATENT APPLICATION: US/10/536,772 TIME: 15:05:39

Input Set : A:\pto.kd.txt
Output Set: N:\CRF4\06132005\J536772.raw

157 gtgcgactgg cacaaggaca aagag 25
159 <210> SEQ ID NO: 4
160 <211> LENGTH: 23
161 <212> TYPE: DNA
162 <213> ORGANISM: Artificial
164 <220> FEATURE:
165 <221> NAME/KEY: misc_feature
166 <222> LOCATION: (1)..(23)
167 <223> OTHER INFORMATION: primer
169 <400> SEQUENCE: 4
170 cgaatgatga cgatccccga gcc 23
172 <210> SEQ ID NO: 5
173 <211> LENGTH: 22
174 <212> TYPE: DNA
175 <213> ORGANISM: Artificial
177 <220> FEATURE:
178 <221> NAME/KEY: misc_feature
179 <222> LOCATION: (1)..(22)
180 <223> OTHER INFORMATION: primer
182 <400> SEQUENCE: 5
183 ccgatgctgc tgacgctggc cg 22
185 <210> SEQ ID NO: 6
186 <211> LENGTH: 25
187 <212> TYPE: DNA
188 <213> ORGANISM: Artificial
190 <220> FEATURE:
191 <221> NAME/KEY: misc_feature
192 <222> LOCATION: (1)..(25)
193 <223> OTHER INFORMATION: primer
195 <400> SEQUENCE: 6
196 tgttggcacc agaaaatcct gcttg 25
198 <210> SEQ ID NO: 7
199 <211> LENGTH: 20
200 <212> TYPE: DNA
201 <213> ORGANISM: Artificial
203 <220> FEATURE:
204 <221> NAME/KEY: misc_feature
205 <222> LOCATION: (1)..(20)
206 <223> OTHER INFORMATION: primer
208 <400> SEQUENCE: 7
209 aagtactccg tgtggatcgg 20
211 <210> SEQ ID NO: 8
212 <211> LENGTH: 20
213 <212> TYPE: DNA
214 <213> ORGANISM: Artificial
216 <220> FEATURE:
217 <221> NAME/KEY: misc_feature
218 <222> LOCATION: (1)..(20)
219 <223> OTHER INFORMATION: primer

RAW SEQUENCE LISTING DATE: 06/13/2005
PATENT APPLICATION: US/10/536,772 TIME: 15:05:39

Input Set : A:\pto.kd.txt
Output Set: N:\CRF4\06132005\J536772.raw

<400> SEQUENCE: 8
tcaagttggg ggacaaaaag 20
<210> SEQ ID NO: 9
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(25)
<223> OTHER INFORMATION: primer
<400> SEQUENCE: 9
gtgcgactgg cacaaggaca aagag 25
<210> SEQ ID NO: 10
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(23)
<223> OTHER INFORMATION: primer
<400> SEQUENCE: 10
ggggatcgtc atcattcgct cct 23
<210> SEQ ID NO: 11
<211> LENGTH: 15
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (1)..(15)
<223> OTHER INFORMATION: oligopeptide corresponding to C-terminus of CT120 protein
<400> SEQUENCE: 11
Cys Arg Lys Ala Val Arg Leu Phe Asp Thr Pro Gln Ala Lys Lys
1 5 10 15
<210> SEQ ID NO: 12
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (1)..(20)
<223> OTHER INFORMATION: primer
<400> SEQUENCE: 12
atgctgctga cgctggccgg 20
<210> SEQ ID NO: 13
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial
<220> FEATURE:
<221> NAME/KEY: misc_feature

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 06/13/2005
PATENT APPLICATION: **US/10/536,772** TIME: 15:05:40

Input Set : **A:\pto.kd.txt**
Output Set: **N:\CRF4\06132005\J536772.raw**

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:3,4,5,6,7,8,9,10,11,12,13

VERIFICATION SUMMARY DATE: 06/13/2005
PATENT APPLICATION: **US/10/536,772** TIME: 15:05:40

Input Set : **A:\pto.kd.txt**
Output Set: **N:\CRF4\06132005\J536772.raw**

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:29 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:27